Unbinned Maximum Likelihood Method for asymmetry extraction at PHENIX

Anselm Vossen[a], John Koster[a]

[a] University of Illinois at Urbana Champaign

Abstract

This writeup describes the unbinned maximum likelihood method to extract physics asymmetries at PHENIX. As a concrete example, the extraction of A_N is considered. A simple approach in which A_N is the only free parameter is described in detail. The corresponding implementation was used to extract A_N from neutral pions in the central arm and a comparison with 'traditional' left-right asymmetries is shown. Several ways to extend this methods are shown which will enable the combination of fills and beams in a natural way. This will allow the use of the full statistics in the future to extract A_N with minimum variance. The method is flexible enough to be extended to asymmetries depending on two azimuthal angles.

Contents

1 Introduction

The maximum likelihood principle is a good method to arrive at a bias free estimator with minimum variance [1]. Since it allows to build an estimator from "‘first principles"’ is also allows for the integration of a priori knowledge which is otherwise difficult to consider. Here an unbinned method is presented, in which each event contributes. This prevents loss of information due to binning. The primary motivation to construct this estimator for PHENIX transverse spin analysis has been to construct a method that is able to deal with time dependent acceptances, more complicated angular dependences and low count rates. Here only an implementation for the most simple case in which the asymmetries are considered fill by fill (and beam by beam) is presented. This has been successfully tested for A_N asymmetries extracted from run8 data in the central arm [3] and results are shown. In addition, several ways to extend this method to integrate fills and beams, have more complicated angular dependences and estimate the observed variance are discussed.

2 Maximum Likelihood Method for A_N

The maximum likelihood estimate of a parameter maximizes the probability of the observation. In the case of the A_N asymmetry the observation are the counts in a specific angular bin. As long as only one particle in the final state is observed, only one azimuthal (around the beam axis) angle (Φ) can be constructed with the spin vector and the momentum of the unpolarized particle. Including acceptance and efficiency Acc and factors like beam flux f, the probability distribution function for events with an angle Φ is

$$P(\Phi_i) = \mathrm{Acc} \cdot f \cdot (1 \pm A_N \cos(\Phi_i))\,. \tag{1}$$

Here we neglect constant factors (like the normalization) because they do not play a role in the minimization. The extended likelihood of the observation of N events is therefore:

$$\mathcal{L}(A_N) = \prod_N \mathrm{Acc} \cdot f\,(1 \pm A_N \cos(\Phi_i))\,\frac{e^{-\mu}\mu^N}{N!}. \tag{2}$$

[1] The sign of A_N depends on the polarization.

The "extension" $\frac{e^{-\mu}\mu^N}{N!}$ is the Poisson-probability to observe the number of events N if the expected number of events is μ.

For now this factor is not important, since a variation of A will not impact μ. However, if for example the efficiency is allowed to change in time, this factor becomes important. Without it, the likelihood could grow infinitely with the choice of an infinite efficiency.

Equation 2 is already instructive how to compute the likelihood.

[1] the likelihood of observing Φ is obviously zero, but (hopefully) w.l.o.g we drop $d\Phi$'s etc. with the same argument that we drop other constant factors

We just look for $\hat{A}_N$ such that $\hat{A}_N = \underset{A_N}{\operatorname{argmax}}\ \mathcal{L}(A_N)$.

For practical reasons one chooses instead to minimize the negative log likelihood. Then the products become sums and the numbers that are computational tractable.

The expected number of events in this basic case is

- $\mu^{\uparrow} = 2\pi * L_{\text{rel}}$
- $\mu^{\downarrow} = 2\pi$

Here L_{rel} is the relative luminosity of polarized protons in the state $\uparrow$ with respect to those in the state $\downarrow$[2]. It is up to each analyzer to determine their own measure of the realitve luminosity.

3 Combining Fills and blue/yellow beam data

In the case of limited statistics, it might be necessary to combine fills and asymmetries in which the yellow or the blue beam was polarized. Otherwise the error estimates become unreliable. The maximum likelihood method is no exception and might develop a bias for small sample sizes. However, it allows to combine asymmetries taken at different times in a more natural way. Just combining statistics might lead to a bias due to time dependent detector efficiencies as discussed in [4]. The solution in this reference is a normalization of the relative luminosity such that $L_{\text{rel}} \approx 1$ by dropping events. However, using a maximum likelihood method, this is not necessary anymore, since for each fill and beam acceptance parameters $a_{\text{fill, beam}}$ can be added, accounting for the possible change over time. These 'nuisance parameters' are estimated together with A_N and might also provide insights into detector behavior over time. As long as only a one dimensional modulation is measured, i.e. only in the angle ϕ, constant acceptance factors are sufficient. The situation is different for correlation measurements. Then the count rates can be differential in two azimuthal angles ϕ_1, ϕ_2. Convolution with the acceptance can then lead to complicated interference effects. Since these effects depend on the respective Fourier coefficients of the acceptance, the acceptance function should have enough degrees of freedom to reflect this. However, in practice the numerical differences between the different orders are so large, that the zeroth coefficient is sufficient. For the case of semi deep inelastic scattering, more details can be found in [5].

4 Estimating the variance of the asymmetry

Currently we estimate the error on the extracted asymmetry as

$$\sigma_{A_N} = - \left. \frac{\partial^2 \ln(\mathcal{L})}{\partial^2 A_N} \right|_{\hat{A}_N} . \tag{3}$$

[2] $\uparrow\downarrow$ have the obvious meaning of polarization states

However, in the future the variance of the asymmetry can be directly estimated from the data. To this end, the likelihood can be written as

$$\mathcal{L}(A_N, \sigma_{A_N}) = \prod_N \mathrm{Acc} \cdot f \left(1 \pm \mathrm{gauss}(A_N, \sigma_{A_N}) \cos(\Phi_i)\right) \frac{e^{-\mu}\mu^N}{N!}. \tag{4}$$

Here both, the asymmetry and its error, are directly estimated from data. This might lead to a more stable estimate and is independent of the statistical models of specific count-rates.

5 Implementation and Results

The Fletcher-Reeves conjugate gradient method as implmented in the GNU scientific library was used [2]. This is a popular example of a conjugate gradient algorithm using line searches to avoid the solution of a quadratic problem in each step. For the solution the minimum is iteratively searched for along a line. The direction of the line is defined by conjugate gradients. At the starting point the gradient of the function is chosen as the direction in which to search. For subsequent steps, the new direction is a conjugate of the old one and the gradient of the function at the new point. This method is only slightly more complicated than usual gradient descent methods, but converges much faster. In this application, where the number of free parameters is very limited, an easier algorithm would also work. In fact, we tested a simplex minimization scheme, that does not require the computation of the derivatives. However, for extensions as outlined in this document, the use of a more advanced algorithm will be necessary. After testing with toy monte-carlo, we tested the unbinned maximum likelihood estimator on the Run-8 dataset to estimate A_N in the central arm. The following figures show the comparison with the 'traditional' left-right asymmetries.

5.1 Code location

The code for the MLE and the toy MC can be found in

`/phenix/u/workarea/vossen/phMLE/`

The integration in the central analysis code in

`/phenix/u/workarea/jkoster4/devel/CentAn/nana/ssa/attic/calcssa_anselm.cc`

Some familiarity with the central arm analysis is assumed here. However, for details on the central arm method, please see [3]. Asymmetries are calculated in the π^0 and η mass windows which include some combinatorial background contribution. The background asymmetry is estimated using the high and low mass windows around the meson mass. Then, using an estimate of the background contamination level, the pure signal asymmetry is extracted. The results shown here are for the un-corrected asymmetries which have yet to be to subtracted. The "traditional" formula for calculating left-right asymmetries in PHENIX is:

$$\epsilon_N = \epsilon_{lumi} = \frac{N_\uparrow - RL \cdot N_\downarrow}{N_\uparrow + RL \cdot N_\downarrow} \tag{5}$$
$$= PA_N \langle cos(\phi) \rangle \tag{6}$$

where P is the polarization and $\langle cos(\phi) \rangle$ is a weighting term for the distribution of particles around the polarization direction and is determined from data. A simple comparison of the results for one particular set of cuts is shown in figure 1.

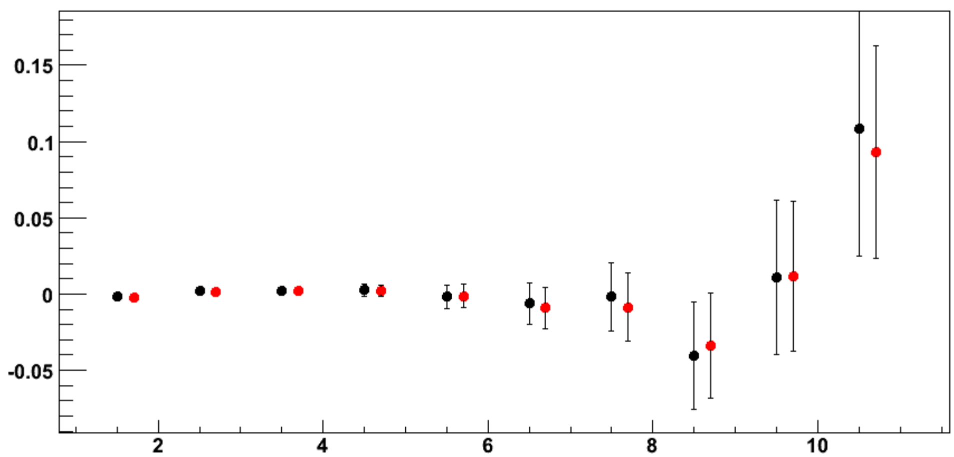

Figure 1: A_N values averaged over all fills of run8pp using even crossings, the blue beam, in the π^0 mass window and plotted against p_T. Black points: traditional luminosity asymmetry calculation, Red points: extended maximum likelihood method.

We aggregated all the asymmetries together to get a sense of their overall agreement. These results are shown in figure 2. In general, the results agree very well. However, for high count rates, there seems to be some deviation. As one would in general expect better agreement for a high number of events, we conclude that the reason is numerical precision. As we use floating point number throughout, the precision gets less, when a large number of events are summed up. This problem can be solved by using higher precision. But since the main application will be analysis with low count rates we did not implement this change yet.

6 Summary

An unbinned estimator for A_N using the maximum likelihood method was introduced. It was implemented and tested on real data. The results were consistent with methods currently used in the analysis. It was shown how to extend the estimator to combine different datasets by introducing nuisance parameters describing possible time dependent acceptances and efficiencies of the detector. This will enable the use of the full

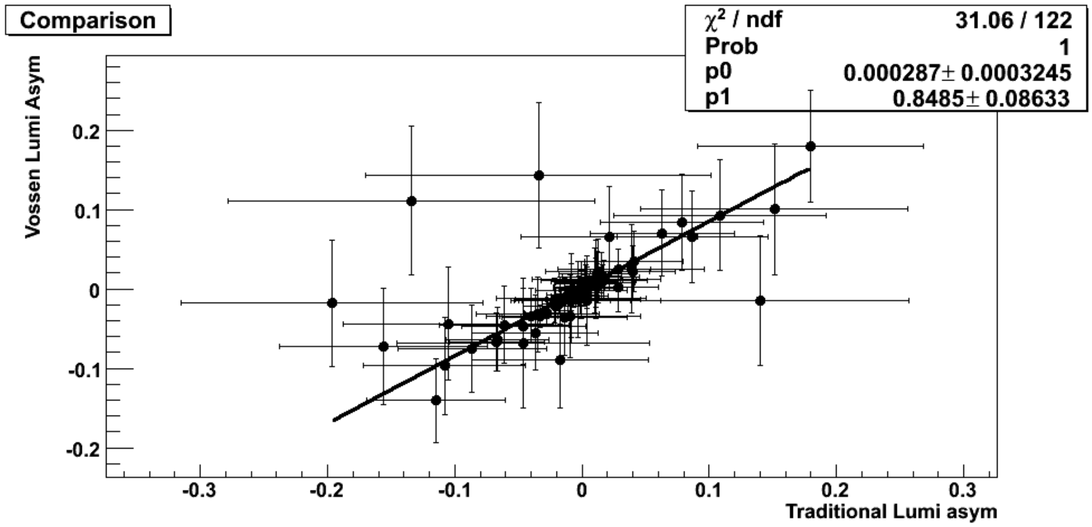

Figure 2: Comparison between the "traditional" and MLE methods. X-axis "traditional" asymmetry values, Y-axis MLE asymmetry values. All four mass bins (π^0 signal, π^0 background, η signal, η background), both beams (blue and yellow), both crossing selections (even and odd) and all p_T points are drawn. The "traditional" asymmetry calculation was tweaked slightly after the plots were made, but still correspond closely to the final results made preliminary on Dec. 11, 2009.

data sample without the need to normalize relative luminosities. The authors plan to implement these in the near future. Together with the natural extension to asymmetries dependent on more than one angle and the possibility to estimate the variance of the observables directly from data, the maximum likelihood estimator described here will be a valuable tool in future, more complex analysis.

References

[1] Frederick James, Statistical Methods in Experimental Physics, 2nd Edition, World Scientific

[2] B. Gough (Editor), GNU Scientific Library Manual, 2nd Edition, Network Theory Ltd., 2006

[3] John Koster et. al, an864

[4] Patricia Liebing, an586

[5] Anselm Vossen, Transverse Spin Asymmetries at the COMPASS experiment, Section 4.6 http://www.freidok.uni-freiburg.de/volltexte/5234/